

Goal

- ❑ Create a natural language machine learning model that can predict the intent and topic of utterances spoken during hospital trauma resuscitations.
- ❑ Intents: Request, Report, Command
- ❑ Topics: Intubation, Blood Pressure, Respiratory Rate...

Motivations and Objectives

- ❑ Motivations
 - Interest in the field of language processing and machine learning.
 - There aren't very many applications of these topics in the medical field.
- ❑ Objectives
 - Classify utterances in two different ways based on intent and topic.
 - Achieve an accuracy greater than 70% for both types of classifications.

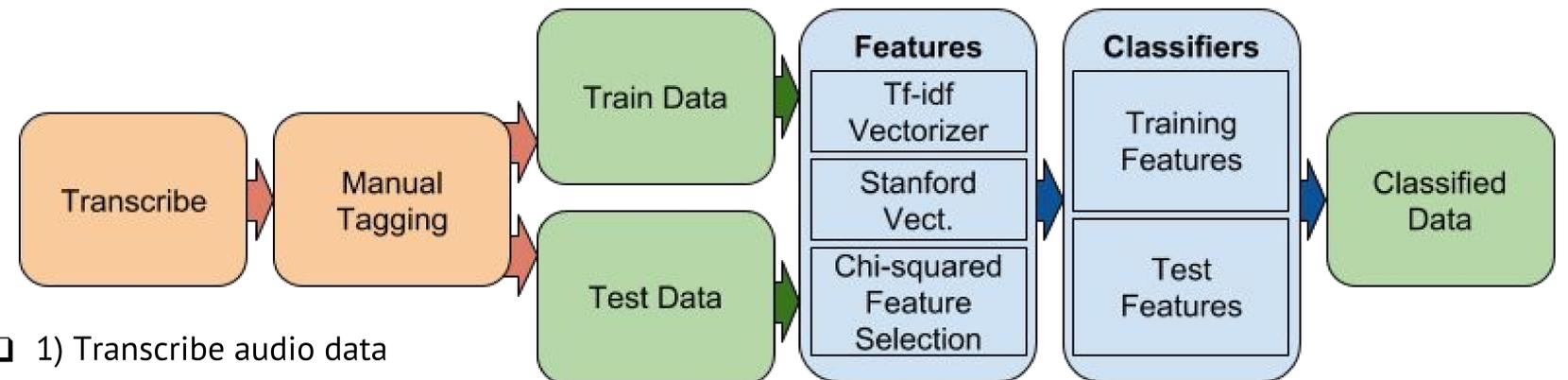
Research Challenges

- ❑ Finding the parameters that yielded the best results for data splitting, feature selection, and normalization.
- ❑ Manually transcribing and classifying large amounts of data.
- ❑ Finding the best classifiers/models that returned the best results.
- ❑ Implementing a non-traditional hierarchical classification model for the topic (second) classifications.

Acknowledgement

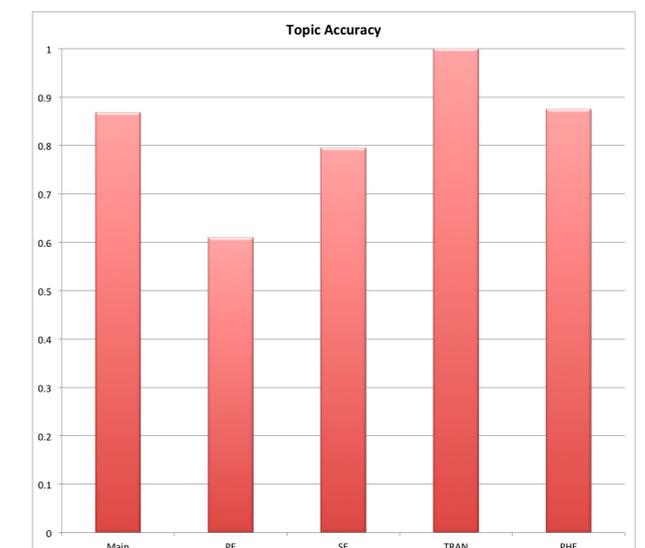
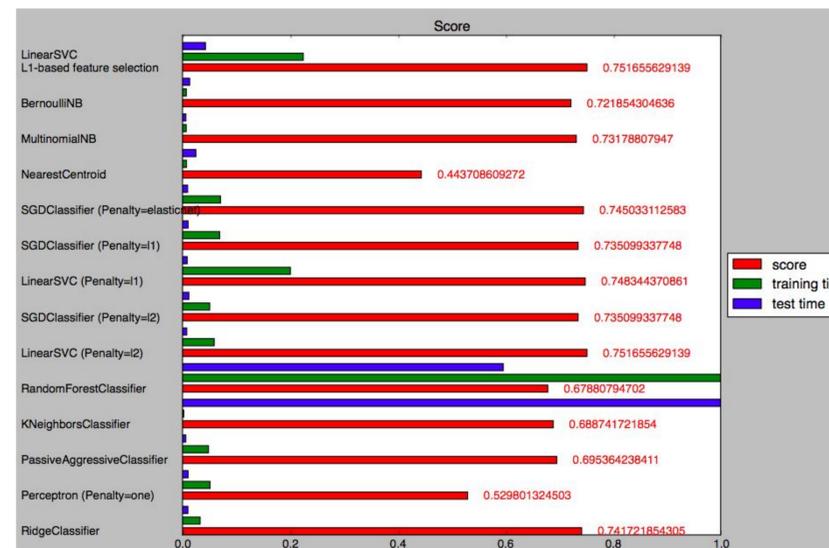
We would like to thank Ivan Marsic and Yue Gu for their guidance and support throughout our project, as well as staff at RWJ for providing us with the data to analyze and making it all possible..

Methodology



- ❑ 1) Transcribe audio data
 - ❑ 2) Manually tag each utterance by intent, topic(main + sub)
 - ❑ 3) Normalize + Split corpus 80/20 (Train/Test)
 - ❑ 4) Create Features, Select Features
 - ❑ 5) Train and Test Classifiers
 - ❑ 6) Measure/ Compare Performance
- ❑ Intent and Topic Classification are done separately
 - ❑ For topic classification, main categories are predicted first then secondary categories

Results



- ❑ Intent Classification
 - Best Accuracy: 75%
 - Classifier: LinearSVC
- ❑ Topic Classification
 - Accuracies left to right: 86%, 61%, 79%, 100%, 86%
 - Classifiers left to right: Ridge, Perceptron, LinearSVC, Ridge, Ridge

References

- [1] TRU-IT Coding Schemes V6
- [2] scikit Learn, <http://scikit-learn.org/stable/index.html>
- [3] Stanford Parser, <http://nlp.stanford.edu/software/lex-parser.shtml>
- [4] Language Disambiguation for Disease Analysis